

Hybridization probe for third leucine rich repeat region

SEQ ID NO:01, nucleotides 1504-2166

Hybridization probe for fourth leucine rich repeat region

SEQ ID NO:01, nucleotides 2167-2751

Hybridization probe for EGF repeats one to five

SEQ ID NO:01, nucleotides 2752-3327

Hybridization probe for the sixth EGF repeat and preceding spacer region

SEQ ID NO:01, nucleotides 3328-3461

Hybridization probe for the 99aa spacer/G-loop region

SEQ ID NO:01, nucleotides 3462-3987

Hybridization probe for EGF repeats seven to nine

SEQ ID NO:01, nucleotides 3988-4341

Hybridization probe for the cysteine knot region

SEQ ID NO:01, nucleotides 4342-4575

Table 6. PCR Primers for regions of Human Slit.

PCR Primers for first leucine rich repeat region

Forward: SEQ ID NO:01, nucleotides 82-111

Reverse: reverse complement of SEQ ID NO:01, nucleotides 799-828

PCR Primers for second leucine rich repeat region

Forward: SEQ ID NO:01, nucleotides 829-858

Reverse: reverse complement of SEQ ID NO:01, nucleotides 1474-1503

PCR Primers for third leucine rich repeat region

Forward: SEQ ID NO:01, nucleotides 1504-1533

Reverse: reverse complement of SEQ ID NO:01, nucleotides 2137-2166

PCR Primers for fourth leucine rich repeat region

Forward: SEQ ID NO:01, nucleotides 2167-2196

Reverse: reverse complement of SEQ ID NO:01, nucleotides 2722-2751

PCR Primers for EGF repeats one to five

Forward: SEQ ID NO:01, nucleotides 2752-2781

Reverse: reverse complement of SEQ ID NO:01, nucleotides 3298-3327

PCR Primers for the sixth EGF repeat and preceding spacer region

Forward: SEQ ID NO:01, nucleotides 3328-3357
Reverse: reverse complement of SEQ ID NO:01, nucleotides 3432-3461

PCR Primers for the 99aa spacer/G-loop region

Forward: SEQ I:01, nucleotides 3462-3491
Reverse: reverse complement of SEQ ID NO:01, nucleotides 3958-3987

PCR Primers for EGF repeats seven to nine

Forward: SEQ ID NO:01, nucleotides 3988-4017
Reverse: reverse complement of SEQ ID NO:01, nucleotides 4312-4341

PCR Primers for the cysteine knot region

Forward: SEQ ID NO:01, nucleotides 4342-4371
Reverse: reverse complement of SEQ ID NO:01, nucleotides 4546-4575

Leucine rich repeats (LRRs) are predicted by comparison with known proteins and by the presence of a leucine rich core sequence. ~~presence of the core sequence: xaaExxLxxLxxLxxLxxLxxLxxL, where x is any amino acid.~~ In slit proteins, the LRRs are flanked by conserved sequences referred to as the amino- and carboxy-flanking regions. These flanking regions are found in other known proteins, but only in a few instances are both the amino- and carboxy- flank regions present in a single protein. ~~The amino flank region is defined by the consensus: CPxxCxC[1-6x]GxxVDCxxxGL[2-4x]αPxxαPxdttx where x is any amino acid, [x] represents a variable number of amino acids and α is a hydrophobic residue. Lower case indicates a residue is not highly conserved at a particular position. The carboxy flank region is defined by the consensus: PβxCγCxxα[1-5x]Wα[14-26x]RCxxPxxxxxxxxαxxxxxxF[1-3x]Cs[3-17x] where β is W or a hydrophobic residue, γ is D or N and α is a hydrophobic residue. Epidermal growth factor (EGF) repeats are predicted by the consensus: CxxxxCxxgxC[6-9x]αxCxCxxGαxGxxCxxxxxx.~~ The so called "99aa spacer" is actually ~200 amino acids in the Drosophila protein and 174 amino acids in Human Slit-1. This region shows homology to the G-loops of laminin A chains.

Cysteine knots are dimerisation domains defined by the presence of six cysteine residues between which disulphide bridges form. The only absolutely conserved residues are the six cysteines, and spacing between them is highly variable, apart from between cysteines 2 and 3, and 5 and 6. ~~C[x]C[1-3x]GxC[x]C[x]CxC~~. The glycine between cysteines 2 and 3 is only present in a subset of cysteine knots. Drosophila slit and Human slit-1 both have an extra cysteine after cysteines 5 and 6: this may serve as an intermolecular bond. Human Slit-1 gene displays the overall structure of the Drosophila gene, and amino acid conservation is found along the entire length of the protein

42

(48% homology at the amino acid sequence excluding the signal sequence; see below). The Human gene has an extra LRR between LRR2 and LRR3 of the first set of LRRs; in the third set, the Human gene has an extra LRR between LRR3 and LRR4. The Human gene has two extra EGF repeats, on either side of the seventh EGF repeat in Drosophila slit.

Isolation of Human slit-1

Searching of the EST database revealed an EST, ab16g10.r1, with homology to the 99aa spacer region of Drosophila slit. This EST was used to probe a Human fetal brain library (Stratagene), and clones for Human slit-1 were isolated.

Features of Human Slit Predicted Protein

| | |
|------------------------------------|--|
| Signal sequence | SEQ ID NO:02, residues 7-24 |
| First amino-flanking sequence | SEQ ID NO:02, residues 28-59 |
| First set of Leucine Rich Repeats | SEQ ID NO:02, residues 60-179 (6 repeats) |
| First carboxy-flanking sequence | SEQ ID NO:02, residues 180-276 |
| Second amino-flanking sequence | SEQ ID NO:02, residues 277-308 |
| Second set of Leucine Rich Repeats | SEQ ID NO:02, residues 309-434 (5 repeats) |
| Second carboxy-flanking sequence | SEQ ID NO:02, residues 435-501 |
| Third amino-flanking sequence | SEQ ID NO:02, residues 502-533 |
| Third set of Leucine Rich Repeats | SEQ ID NO:02, residues 534-560 (5 repeats) |
| Third carboxy-flanking sequence | SEQ ID NO:02, residues 661-722 |
| Fourth amino-flanking sequence | SEQ ID NO:02, residues 723-754 |
| Fourth set of Leucine Rich Repeats | SEQ ID NO:02, residues 755-855 (4 repeats) |
| Fourth carboxy-flanking sequence | SEQ ID NO:02, residues 856-917 |
| First EGF repeat | SEQ ID NO:02, residues 918-952 |
| Second EGF repeat | SEQ ID NO:02, residues 953-993 |
| Third EGF repeat | SEQ ID NO:02, residues 994-1031 |
| Fourth EGF repeat | SEQ ID NO:02, residues 1032-1071 |
| Fifth EGF repeat | SEQ ID NO:02, residues 1072-1109 |
| Spacer | SEQ ID NO:02, residues 1110-1116 |

| | |
|---------------------|----------------------------------|
| Sixth EGF repeat | SEQ ID NO:02, residues 1117-1153 |
| “99aa spacer” | SEQ ID NO:02, residues 1155-1329 |
| Seventh EGF repeat | SEQ ID NO:02, residues 1330-1366 |
| Eighth EGF repeat | SEQ ID NO:02, residues 1367-1404 |
| Nineth EGF repeat | SEQ ID NO:02, residues 1405-1447 |
| Cysteine knot motif | SEQ ID NO:02, residues 1448-1525 |

Amino acid identity between Drosophila and Human Slit-1

| | |
|------------------------------------|-----------------------------------|
| First amino-flanking sequence | 53% |
| First set of Leucine Rich Repeats | 52% (54%, 67%, NA, 38%, 54%, 50%) |
| First carboxy-flanking sequence | 42% |
| Second amino-flanking sequence | 50% |
| Second set of Leucine Rich Repeats | 60% (54%, 58%, 67%, 71%, 50%) |
| Second carboxy-flanking sequence | 62% |
| Third amino-flanking sequence | 56% |
| Third set of Leucine Rich Repeats | 49% (46%, 46%, 42%, NA, 58%) |
| Third carboxy-flanking sequence | 36% |
| Fourth amino-flanking sequence | 53% |
| Fourth set of Leucine Rich Repeats | 48% (25%, 58%, 46%, 63%) |
| Fourth carboxy-flanking sequence | 63% |
| First EGF repeat | 34% |
| Second EGF repeat | 46% |
| Third EGF repeat | 46% |
| Fourth EGF repeat | 35% |
| Fifth EGF repeat | 47% |
| Spacer | 22% |
| Sixth EGF repeat | 40% |
| “99aa spacer” | 38% |